

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/854,774

DATE: 12/10/2001  
TIME: 16:24:18

Input Set : N:\paola\09854774.txt  
Output Set: N:\CRF3\12102001\I854774.raw

## SEQUENCE LISTING

## 4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: FROMMER, Wolf-Bernd  
8 (ii) TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID  
9 TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS  
10 CONTAINING A TRANSPORTER AND THEIR USE

12 (iii) NUMBER OF SEQUENCES: 4

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen  
16 (B) STREET: 1180 Avenue of the Americas  
17 (C) CITY: New York  
18 (D) STATE: NY  
19 (E) COUNTRY: US  
20 (F) ZIP: 10036-8403

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk  
24 (B) COMPUTER: IBM PC compatible  
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/854,774  
C--> 30 (B) FILING DATE: 14-May-2001  
31 (C) CLASSIFICATION: 800

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/362,512  
35 (B) FILING DATE: 1992-07-15

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Meilman, Edward A.  
39 (B) REGISTRATION NUMBER: 24,735  
40 (C) REFERENCE/DOCKET NUMBER: P/951-107

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: (212) 382-0700  
44 (B) TELEFAX: (212) 382-0888  
45 (C) TELEX: 236925

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 1685 base pairs  
52 (B) TYPE: nucleic acid  
53 (C) STRANDEDNESS: single  
54 (D) TOPOLOGY: linear

56 (ii) MOLECULE TYPE: cDNA

58 (vi) ORIGINAL SOURCE:

59 (A) ORGANISM: Arabidopsis thaliano

61 (ix) FEATURE:

62 (A) NAME/KEY: CDS  
63 (B) LOCATION: 57..1511  
64 (D) OTHER INFORMATION: /note= "amino acid transporter"

ENTERED

TECH CENTER 1600/2900

JAN 18 2002

RECEIVED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/854,774

DATE: 12/10/2001

TIME: 16:24:18

Input Set : N:\paola\09854774.txt

Output Set: N:\CRF3\12102001\I854774.raw

```

67      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
69 CTTAAACAT TTATTTTATC TTCTTCTTGT TCTCTCTTTC TCTTTCTCTC ATCACT      56
71 ATG AAG AGT TTC AAC ACA GAA GGA CAC AAC CAC TCC ACG GCG GAA TCC      104
72 Met Lys Ser Phe Asn Thr Glu Gly His Asn His Ser Thr Ala Glu Ser
73   1           5           10           15
75 GGC GAT GCC TAC ACC GTG TCG GAC CCG ACA AAG AAC GTC GAT GAA GAT      152
76 Gly Asp Ala Tyr Thr Val Ser Asp Pro Thr Lys Asn Val Asp Glu Asp
77           20           25           30
79 GGT CGA GAG AAG CGT ACC GGG ACG TGG CTT ACG GCG AGT GCG CAT ATT      200
80 Gly Arg Glu Lys Arg Thr Gly Thr Trp Leu Thr Ala Ser Ala His Ile
81           35           40           45
83 ATC ACG GCG GTG ATA GGC TCC GGA GTG TTG TCT TTA GCA TGG GCT ATA      248
84 Ile Thr Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile
85   50           55           60
87 GCT CAG CTT GGT TGG ATC GCA GGG ACA TCG ATC TTA CTC ATT TTC TCG      296
88 Ala Gln Leu Gly Trp Ile Ala Gly Thr Ser Ile Leu Leu Ile Phe Ser
89 65           70           75           80
91 TTC ATT ACT TAC TTC ACC TCC ACC ATG CTT GCC GAT TGC TAC CGT GCG      344
92 Phe Ile Thr Tyr Phe Thr Ser Thr Met Leu Ala Asp Cys Tyr Arg Ala
93           85           90           95
95 CCG GAT CCC GTC ACC GGA AAA CGG AAT TAC ACT TAC ATG GAC GTT GTT      392
96 Pro Asp Pro Val Thr Gly Lys Arg Asn Tyr Thr Tyr Met Asp Val Val
97   100           105           110
99 CGA TCT TAC CTC GGT GGT AGG AAA GTG CAG CTC TGT GGA GTG GCA CAA      440
100 Arg Ser Tyr Leu Gly Gly Arg Lys Val Gln Leu Cys Gly Val Ala Gln
101   115           120           125
103 TAT GGG AAT CTG ATT GGG GTC ACT GTT GGT TAC ACC ATC ACT GCT TCT      488
104 Tyr Gly Asn Leu Ile Gly Val Thr Val Gly Tyr Thr Ile Thr Ala Ser
105   130           135           140
107 ATT AGT TTG GTA GCG GTA GGG AAA TCG AAC TGC TTC CAC GAT AAA GGG      536
108 Ile Ser Leu Val Ala Val Gly Lys Ser Asn Cys Phe His Asp Lys Gly
109 145           150           155           160
111 CAC ACT GCG GAT TGT ACT ATA TCG AAT TAT CCG TAT ATG GCG GTT TTT      584
112 His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe
113           165           170           175
115 GGT ATC ATT CAA GTT ATT CTT AGC CAG ATC CCA AAT TTC CAC AAG CTC      632
116 Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu
117           180           185           190
119 TCT TTT CTT TCC ATT ATG GCC GCA GTC ATG TCC TTT ACT TAT GCA ACT      680
120 Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr
121   195           200           205
123 ATT GGA ATC GGT CTA GCC ATC GCA ACC GTC GCA GGT GGG AAA GTG GGT      728
124 Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly
125   210           215           220
127 AAG ACG AGT ATG ACG GGC ACA GCG GTT GGA GTA GAT GTA ACC GCA GCT      776
128 Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala
129 225           230           235           240
131 CAA AAG ATA TGG AGA TCG TTT CAA GCG GTT GGG GAC ATA GCG TTC GCC      824
132 Gln Lys Ile Trp Arg Ser Phe Gln Ala Val Gly Asp Ile Ala Phe Ala

```

## RAW SEQUENCE LISTING

DATE: 12/10/2001

PATENT APPLICATION: US/09/854,774

TIME: 16:24:18

Input Set : N:\paola\09854774.txt

Output Set: N:\CRF3\12102001\I854774.raw

133		245		250		255		
135	TAT GCT TAT GCC ACG GTT CTC ATC GAG ATT CAG GAT ACA CTA AGA TCT							872
136	Tyr Ala Tyr Ala Thr Val Leu Ile Glu Ile Gln Asp Thr Leu Arg Ser							
137		260		265		270		
139	AGC CCA GCT GAG AAC AAA GCC ATG AAA AGA GCA AGT CTT GTG GGA GTA							920
140	Ser Pro Ala Glu Asn Lys Ala Met Lys Arg Ala Ser Leu Val Gly Val							
141		275		280		285		
143	TCA ACC ACC ACT TTT TTC TAC ATC TTA TGT GGA TGC ATC GGC TAT GCT							968
144	Ser Thr Thr Thr Phe Phe Tyr Ile Leu Cys Gly Cys Ile Gly Tyr Ala							
145		290		295		300		
147	GCA TTT GGA AAC AAT GCC CCT GGA GAT TTC CTC ACA GAT TTC GGG TTT							1016
148	Ala Phe Gly Asn Asn Ala Pro Gly Asp Phe Leu Thr Asp Phe Gly Phe							
149	305		310		315		320	
151	TTC GAG CCC TTT TGG CTC ATT GAC TTT GCA AAC GCT TGC ATC GCT GTC							1064
152	Phe Glu Pro Phe Trp Leu Ile Asp Phe Ala Asn Ala Cys Ile Ala Val							
153		325		330		335		
155	CAC CTT ATT GGT GCC TAT CAG GTG TTC GCG CAG CCG ATA TTC CAG TTT							1112
156	His Leu Ile Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Gln Phe							
157		340		345		350		
159	GTT GAG AAA AAA TGC AAC AGA AAC TAT CCA GAC AAC AAG TTC ATC ACT							1160
160	Val Glu Lys Lys Cys Asn Arg Asn Tyr Pro Asp Asn Lys Phe Ile Thr							
161		355		360		365		
163	TCT GAA TAT TCA GTA AAC GTA CCT TTC CTT GGA AAA TTC AAC ATT AGC							1208
164	Ser Glu Tyr Ser Val Asn Val Pro Phe Leu Gly Lys Phe Asn Ile Ser							
165		370		375		380		
167	CTC TTC AGA TTG GTG TGG AGG ACA GCT TAT GTG GTT ATA ACC ACT GTT							1256
168	Leu Phe Arg Leu Val Trp Arg Thr Ala Tyr Val Ile Thr Thr Val							
169	385		390		395		400	
171	GTA GCT ATG ATA TTC CCT TTC TTC AAC GCG ATC TTA GGT CTT ATC GGA							1304
172	Val Ala Met Ile Phe Pro Phe Phe Asn Ala Ile Leu Gly Leu Ile Gly							
173		405		410		415		
175	GCA GCT TCC TTC TGG CCT TTA ACG GTT TAT TTC CCT GTG GAG ATG CAC							1352
176	Ala Ala Ser Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met His							
177		420		425		430		
179	ATT GCA CAA ACC AAG ATT AAG AAG TAC TCT GCT AGA TGG ATT GCG CTG							1400
180	Ile Ala Gln Thr Lys Ile Lys Lys Tyr Ser Ala Arg Trp Ile Ala Leu							
181		435		440		445		
183	AAA ACG ATG TGC TAT GTT TGC TTG ATC GTC TCG CTC TTA GCT GCA GCC							1448
184	Lys Thr Met Cys Tyr Val Cys Leu Ile Val Ser Leu Leu Ala Ala Ala							
185		450		455		460		
187	GGA TCC ATC GCA GGA CTT ATA AGT AGT GTC AAA ACC TAC AAG CCC TTC							1496
188	Gly Ser Ile Ala Gly Leu Ile Ser Ser Val Lys Thr Tyr Lys Pro Phe							
189	465		470		475		480	
191	CGG ACT ATG CAT GAG TGAGTTTGAG ATCCTCAAGA GAGTCAAAAA TATATGTAGT							1551
192	Arg Thr Met His Glu							
193		485						
195	AGTTTGGTCT TTCTGTTAAA CTATCTGGTG TCTAAATCCA ATGAGAATGC TTTATTGCTA							1611
197	AAACTTCATG AATCTCTCTG TATCTACATC TTTCAATCTA ATACATATGA GCTCTTCCAA							1671
199	AAAAAAAAAAAA AAAA							1685

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/854,774

DATE: 12/10/2001

TIME: 16:24:18

Input Set : N:\paola\09854774.txt

Output Set: N:\CRF3\12102001\I854774.raw

202 (2) INFORMATION FOR SEQ ID NO: 2:

204 (i) SEQUENCE CHARACTERISTICS:

205 (A) LENGTH: 485 amino acids

206 (B) TYPE: amino acid

207 (D) TOPOLOGY: linear

209 (ii) MOLECULE TYPE: protein

211 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

213 Met Lys Ser Phe Asn Thr Glu Gly His Asn His Ser Thr Ala Glu Ser
214 1      5      10      15
216 Gly Asp Ala Tyr Thr Val Ser Asp Pro Thr Lys Asn Val Asp Glu Asp
217      20      25      30
219 Gly Arg Glu Lys Arg Thr Gly Thr Trp Leu Thr Ala Ser Ala His Ile
220      35      40      45
222 Ile Thr Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile
223      50      55      60
225 Ala Gln Leu Gly Trp Ile Ala Gly Thr Ser Ile Leu Leu Ile Phe Ser
226      65      70      75      80
228 Phe Ile Thr Tyr Phe Thr Ser Thr Met Leu Ala Asp Cys Tyr Arg Ala
229      85      90      95
231 Pro Asp Pro Val Thr Gly Lys Arg Asn Tyr Thr Tyr Met Asp Val Val
232      100     105     110
234 Arg Ser Tyr Leu Gly Gly Arg Lys Val Gln Leu Cys Gly Val Ala Gln
235      115     120     125
237 Tyr Gly Asn Leu Ile Gly Val Thr Val Gly Tyr Thr Ile Thr Ala Ser
238      130     135     140
240 Ile Ser Leu Val Ala Val Gly Lys Ser Asn Cys Phe His Asp Lys Gly
241 145     150     155     160
243 His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe
244      165     170     175
246 Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu
247      180     185     190
249 Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr
250      195     200     205
252 Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly
253      210     215     220
255 Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala
256 225     230     235     240
258 Gln Lys Ile Trp Arg Ser Phe Gln Ala Val Gly Asp Ile Ala Phe Ala
259      245     250     255
261 Tyr Ala Tyr Ala Thr Val Leu Ile Glu Ile Gln Asp Thr Leu Arg Ser
262      260     265     270
264 Ser Pro Ala Glu Asn Lys Ala Met Lys Arg Ala Ser Leu Val Gly Val
265      275     280     285
267 Ser Thr Thr Thr Phe Phe Tyr Ile Leu Cys Gly Cys Ile Gly Tyr Ala
268      290     295     300
270 Ala Phe Gly Asn Asn Ala Pro Gly Asp Phe Leu Thr Asp Phe Gly Phe
271 305     310     315     320
273 Phe Glu Pro Phe Trp Leu Ile Asp Phe Ala Asn Ala Cys Ile Ala Val
274      325     330     335

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/854,774

DATE: 12/10/2001

TIME: 16:24:18

Input Set : N:\paola\09854774.txt

Output Set: N:\CRF3\12102001\I854774.raw

```

276 His Leu Ile Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Gln Phe
277          340          345          350
279 Val Glu Lys Lys Cys Asn Arg Asn Tyr Pro Asp Asn Lys Phe Ile Thr
280          355          360          365
282 Ser Glu Tyr Ser Val Asn Val Pro Phe Leu Gly Lys Phe Asn Ile Ser
283          370          375          380
285 Leu Phe Arg Leu Val Trp Arg Thr Ala Tyr Val Val Ile Thr Thr Val
286 385          390          395          400
288 Val Ala Met Ile Phe Pro Phe Phe Asn Ala Ile Leu Gly Leu Ile Gly
289          405          410          415
291 Ala Ala Ser Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met His
292          420          425          430
294 Ile Ala Gln Thr Lys Ile Lys Lys Tyr Ser Ala Arg Trp Ile Ala Leu
295          435          440          445
297 Lys Thr Met Cys Tyr Val Cys Leu Ile Val Ser Leu Leu Ala Ala Ala
298          450          455          460
300 Gly Ser Ile Ala Gly Leu Ile Ser Ser Val Lys Thr Tyr Lys Pro Phe
301 465          470          475          480
303 Arg Thr Met His Glu
304          485

```

306 (2) INFORMATION FOR SEQ ID NO: 3:

308 (i) SEQUENCE CHARACTERISTICS:

309 (A) LENGTH: 1740 base pairs

310 (B) TYPE: nucleic acid

311 (C) STRANDEDNESS: single

312 (D) TOPOLOGY: linear

314 (ii) MOLECULE TYPE: cDNA

316 (vi) ORIGINAL SOURCE:

317 (A) ORGANISM: Arabidopsis thaliana

319 (ix) FEATURE:

320 (A) NAME/KEY: CDS

321 (B) LOCATION: 80..1558

322 (D) OTHER INFORMATION: /product= "amino acid transporter"

325 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

327 CTATTTTATA ATCCTCTTC TTTTGTTCA TAGCTTTGTA ATTATAGTCT TATTTCTCTT      60
329 TAAGGCTCAA TAAGAGGAG ATG GGT GAA ACC GCT GCC GCC AAT AAC CAC CGT      112
330          Met Gly Glu Thr Ala Ala Ala Asn Asn His Arg
331          1          5          10
333 CAC CAC CAC CAT CAC GGC CAC CAG GTC TTT GAC GTG GCC AGC CAC GAT      160
334 His His His His His Gly His Gln Val Phe Asp Val Ala Ser His Asp
335          15          20          25
337 TTC GTC CCT CCA CAA CCG GCT TTT AAA TGC TTC GAT GAT GAT GGC CGC      208
338 Phe Val Pro Pro Gln Pro Ala Phe Lys Cys Phe Asp Asp Asp Gly Arg
339          30          35          40
341 CTC AAA AGA ACT GGG ACT GTT TGG ACC GCG AGC GCT CAT ATA ATA ACT      256
342 Leu Lys Arg Thr Gly Thr Val Trp Thr Ala Ser Ala His Ile Ile Thr
343          45          50          55
345 GCG GTT ATC GGA TCC GGC GTT TTG TCA TTG GCG TGG GCG ATT GCA CAG      304
346 Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile Ala Gln

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/854,774

DATE: 12/10/2001

TIME: 16:24:19

Input Set : N:\paola\09854774.txt

Output Set: N:\CRF3\12102001\I854774.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]